

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT:

Seidel, Christoph; Weinhues, Ursula-Henrike;  
Schmitt, Urban; Motz, Manfred; Wiedmann, Michael;  
Upmeier, Barbara; Soutscheck, Erwin

(ii) TITLE OF INVENTION:

Recombinant antigen from the NS3 region of the hepatitis C virus

(iii) NUMBER OF SEQUENCES: 9

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Felte & Lynch  
(B) STREET: 805 Third Avenue  
(C) CITY: New York  
(D) STATE: New York  
(E) COUNTRY: USA  
(F) ZIP: 10022

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette  
(B) COMPUTER: IBM PS/2  
(C) OPERATING SYSTEM: PC-DOS  
(D) SOFTWARE: PatentIn Release #1.0,  
Version #1.25 (EPA)

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:  
(B) FILING DATE:  
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 08/511,759  
(B) FILING DATE: 7-AUGUST-1995y

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: p 44 28 705.4  
(B) FILING DATE: 12 AUGUST 1994

(viii) ATTORNEY/AGENT INFORMATION

(A) NAME: HANSON, NORMAN  
(B) REGISTRATION NUMBER: 30, 946  
(C) REFERENCE/DOCKET NUMBER: HUBR 1067.1

(ix) TELECOMMUNICATION INFORMATION

(A) TELEPHONE: 212-688-9200  
(B) TELEFAX: 212-838-3884

106290-22096860

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 885 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) INITIAL ORIGIN:

- (A) ORGANISM: hepatitis C virus

(viii) POSITION IN GENOME:

- (A) CHROMOSOME SEGMENT: NS3

(ix) CHARACTERISTICS:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..885

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ATG ACC ATG ATT ACG AAT TCC CGG GGA TCC ATC ATG AAA TCC CCG GTG 48  
Met Thr Met Ile Thr Asn Ser Arg Gly Ser Ile Met Lys Ser Pro Val  
1 5 10 15

TTC ACG GAT AAC TCC TCT CCA CCG GTA GTG CCC CAG AGC TTC CAG GTG 96  
Phe Thr Asp Asn Ser Ser Pro Pro Val Val Pro Gln Ser Phe Gln Val  
20 25 30

GCT CAC CTG CAT GCT CCC ACA GGC AGC GGC AAG AGC ACC AAG GTC CCG 144  
Ala His Leu His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys Val Pro  
35 40 45

GCT GCA TAC GCA GCT CAG GGC TAC AAG GTG CTA GTG CTC AAC CCT TCT 192  
Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro Ser  
50 55 60

GTT GCT GCA ACA TTG GGC TTT GGT GCC TAC ATG TCC AAG GCT CAT GGG 240  
Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His Gly  
65 70 75 80

ATC GAT CCT AAC ATC AGG ACC GGG GTG AGA ACA ATT ACC ACT GGC AGC 288  
Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly Ser  
85 90 95

CCC ATT ACG TAC TCC ACT TAC GGC AAG TTT CTT GCC GAC GGC GGG TGC 336  
Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly Cys  
100 105 110

GCA	GGG	GGT	GCT	TAT	GAC	ATA	ATA	ATT	TGT	GAC	GAG	TGC	CAC	TCC	ACG	384
Ala	Gly	Gly	Ala	Tyr	Asp	Ile	Ile	Ile	Cys	Asp	Glu	Cys	His	Ser	Thr	
	115					120					125					
GAT	GCC	ACA	TCC	ATC	TTG	GGC	ATC	GGC	ACT	GTC	CTT	GAC	CAA	GGA	GAG	432
Asp	Ala	Thr	Ser	Ile	Leu	Gly	Ile	Gly	Thr	Val	Leu	Asp	Gln	Gly	Glu	
	130					135					140					
ACT	GCG	GGG	GCG	AAA	TTG	GTT	GTG	TTC	GCC	ACC	GCC	ACC	CCT	CCG	GGC	480
Thr	Ala	Gly	Ala	Lys	Leu	Val	Val	Phe	Ala	Thr	Ala	Thr	Pro	Pro	Gly	
	145				150				155						160	
TCC	GTC	ACT	GTG	CCC	CAT	CCC	AAC	ATT	GAG	GAG	GTT	GCT	CTA	TCC	ACC	528
Ser	Val	Thr	Val	Pro	His	Pro	Asn	Ile	Glu	Glu	Val	Ala	Leu	Ser	Thr	
				165					170					175		
ACC	GGA	GAG	ATC	CCT	TTT	TAC	GGC	AAG	GCT	ATC	CCC	CTT	GAG	GTA	ATC	576
Thr	Gly	Glu	Ile	Pro	Phe	Tyr	Gly	Lys	Ala	Ile	Pro	Leu	Glu	Val	Ile	
			180					185					190			
AAG	GGG	GGG	AGA	CAT	CTC	ATC	TTC	TGT	CAT	TCA	AAG	AGG	AAG	TGC	GAT	624
Lys	Gly	Gly	Arg	His	Leu	Ile	Phe	Cys	His	Ser	Lys	Arg	Lys	Cys	Asp	
		195					200					205				
GAG	CTC	GCC	ACA	AAG	CTG	GTC	GCA	ATG	GGC	ATC	AAT	GCC	GTG	GCC	TAC	672
Glu	Leu	Ala	Thr	Lys	Leu	Val	Ala	Met	Gly	Ile	Asn	Ala	Val	Ala	Tyr	
	210					215					220					
TAC	CGC	GGT	CTT	GAC	GTG	TCC	GTC	ATC	CCG	ACC	AGC	GGT	GAT	GTT	GTC	720
Tyr	Arg	Gly	Leu	Asp	Val	Ser	Val	Ile	Pro	Thr	Ser	Gly	Asp	Val	Val	
	225				230				235						240	
GTC	GTG	GCA	ACC	GAC	GCC	CTC	ATG	ACC	GGC	TAT	ACC	GGC	GAC	TTC	GAC	768
Val	Val	Ala	Thr	Asp	Ala	Leu	Met	Thr	Gly	Tyr	Thr	Gly	Asp	Phe	Asp	
				245					250					255		
TCG	GTG	ATA	GAC	TGC	AAC	ACG	TGT	GTC	ACT	CAG	ACA	GTC	GAT	TTC	AGC	816
Ser	Val	Ile	Asp	Cys	Asn	Thr	Cys	Val	Thr	Gln	Thr	Val	Asp	Phe	Ser	
			260					265					270			
CTT	GAC	CCT	ACC	TTC	ACC	ATT	GAG	ACG	ACC	ACA	CTT	CCC	CAG	GAT	GCT	864
Leu	Asp	Pro	Thr	Phe	Thr	Ile	Glu	Thr	Thr	Thr	Leu	Pro	Gln	Asp	Ala	
		275					280					285				
GTC	TCC	CGC	ACT	CAA	CGA	CGG										885
Val	Ser	Arg	Thr	Gln	Arg	Arg										
	290					295										

(3) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 295 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Thr Met Ile Thr Asn Ser Arg Gly Ser Ile Met Lys Ser Pro Val  
1 5 10 15  
Phe Thr Asp Asn Ser Ser Pro Pro Val Val Pro Gln Ser Phe Gln Val  
20 25 30  
Ala His Leu His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys Val Pro  
35 40 45  
Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro Ser  
50 55 60  
Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His Gly  
65 70 75 80  
Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly Ser  
85 90 95  
Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly Cys  
100 105 110  
Ala Gly Gly Ala Tyr Asp Ile Ile Ile Cys Asp Glu Cys His Ser Thr  
115 120 125  
Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr Val Leu Asp Gln Gly Glu  
130 135 140  
Thr Ala Gly Ala Lys Leu Val Val Phe Ala Thr Ala Thr Pro Pro Gly  
145 150 155 160  
Ser Val Thr Val Pro His Pro Asn Ile Glu Glu Val Ala Leu Ser Thr  
165 170 175  
Thr Gly Glu Ile Pro Phe Tyr Gly Lys Ala Ile Pro Leu Glu Val Ile  
180 185 190  
Lys Gly Gly Arg His Leu Ile Phe Cys His Ser Lys Arg Lys Cys Asp  
195 200 205

Glu Leu Ala Thr Lys Leu Val Ala Met Gly Ile Asn Ala Val Ala Tyr  
 210 215 220  
 Tyr Arg Gly Leu Asp Val Ser Val Ile Pro Thr Ser Gly Asp Val Val  
 225 230 235 240  
 Val Val Ala Thr Asp Ala Leu Met Thr Gly Tyr Thr Gly Asp Phe Asp  
 245 250 255  
 Ser Val Ile Asp Cys Asn Thr Cys Val Thr Gln Thr Val Asp Phe Ser  
 260 265 270  
 Leu Asp Pro Thr Phe Thr Ile Glu Thr Thr Thr Leu Pro Gln Asp Ala  
 275 280 285  
 Val Ser Arg Thr Gln Arg Arg  
 290 295

(4) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 40 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

AAGGGATCCA TCATGAAATC CCCGGTGTTC ACGGATAACT

40

(5) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 39 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GGGAAGCCTT AATTCTTACC GTCGTTGAGT GCGGGAGAC

[illegible]

## (ii) MOLECULE TYPE: cDNA

GAGGGATCCA TCATGAAAGC GGTGGACTTT ATCCCTGTG

39

(ii) MOLECULE TYPE: cDNA

GAGAAGCTTT TAACACGTGT TGCAGTCTAT CAC

33

## (ii) MOLECULE TYPE: cDNA

GAGGGATCCA TCATGAAACA CCTGCATGCT CCCACCGGC

39

(8) INFORMATION FOR SEQ ID NO: 8:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 33 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GAGAAGCTTT TAATACCAAG CACAGCCTGC GTC

33

(3) INFORMATION FOR SEQ ID NO: 9:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 302 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

Pro Val Glu Asn Leu Glu Thr Thr Met Arg Ser Pro Val Phe Thr  
1 5 10 15  
Asp Asn Ser Ser Pro Pro Val Val Pro Gln Ser Phe Gln Val Ala  
20 25 30  
His Leu His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys Val Pro  
35 40 45  
Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro  
50 55 60  
Ser Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala  
65 70 75  
His Gly Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr  
80 85 90  
Thr Gly Ser Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala  
95 100 105

13.

assum 2

This is  
14-forward  
in 2

[illegible]



SEQUENCE PROTOCOL

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: Boehringer Mannheim GmbH
- (B) ROAD: Sandhofer Str. 112-132
- (C) CITY: Mannheim
- (E) COUNTRY: Germany
- (F) POSTAL CODE: 68305

(ii) TITLE OF APPLICATION: Recombinant antigen from  
the NS3 region of the hepatitis C virus

(iii) NUMBER OF SEQUENCES: 8

(iv) COMPUTER READABLE FORM:

- (A) DATA CARRIER: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0,  
Version #1.25 (EPA)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 885 base pairs
- (B) TYPE: nucleic acid
- (C) STRAND FORM: both
- (D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: cDNA

FOR "22096360"

(vi) INITIAL ORIGIN:

(A) ORGANISM: hepatitis C virus

(viii) POSITION IN THE GENOME:

(A) CHROMOSOME/SEGMENT: NS3

(ix) CHARACTERISTICS:

(A) NAME/KEY: CDS

(B) LOCATION: 1..885

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ATG	ACC	ATG	ATT	ACG	AAT	TCC	CGG	GGA	TCC	ATC	ATG	AAA	TCC	CCG	GTG	48
Met	Thr	Met	Ile	Thr	Asn	Ser	Arg	Gly	Ser	Ile	Met	Lys	Ser	Pro	Val	
1				5				10						15		
TTC	ACG	GAT	AAC	TCC	TCT	CCA	CCG	GTA	GTG	CCC	CAG	AGC	TTC	CAG	GTG	96
Phe	Thr	Asp	Asn	Ser	Ser	Pro	Pro	Val	Val	Pro	Gln	Ser	Phe	Gln	Val	
			20					25					30			
GCT	CAC	CTG	CAT	GCT	CCC	ACA	GGC	AGC	GGC	AAG	AGC	ACC	AAG	GTC	CCG	144
Ala	His	Leu	His	Ala	Pro	Thr	Gly	Ser	Gly	Lys	Ser	Thr	Lys	Val	Pro	
		35					40					45				
GCT	GCA	TAC	GCA	GCT	CAG	GGC	TAC	AAG	GTG	CTA	GTG	CTC	AAC	CCT	TCT	192
Ala	Ala	Tyr	Ala	Ala	Gln	Gly	Tyr	Lys	Val	Leu	Val	Leu	Asn	Pro	Ser	
	50					55					60					
GTT	GCT	GCA	ACA	TTG	GGC	TTT	GGT	GCC	TAC	ATG	TCC	AAG	GCT	CAT	GGG	240
Val	Ala	Ala	Thr	Leu	Gly	Phe	Gly	Ala	Tyr	Met	Ser	Lys	Ala	His	Gly	
65					70					75					80	
ATC	GAT	CCT	AAC	ATC	AGG	ACC	GGG	GTG	AGA	ACA	ATT	ACC	ACT	GGC	AGC	288
Ile	Asp	Pro	Asn	Ile	Arg	Thr	Gly	Val	Arg	Thr	Ile	Thr	Thr	Gly	Ser	
				85					90					95		
CCC	ATT	ACG	TAC	TCC	ACT	TAC	GGC	AAG	TTT	CTT	GCC	GAC	GGC	GGG	TGC	336
Pro	Ile	Thr	Tyr	Ser	Thr	Tyr	Gly	Lys	Phe	Leu	Ala	Asp	Gly	Gly	Cys	
			100					105					110			
GCA	GGG	GGT	GCT	TAT	GAC	ATA	ATA	ATT	TGT	GAC	GAG	TGC	CAC	TCC	ACG	384
Ala	Gly	Gly	Ala	Tyr	Asp	Ile	Ile	Ile	Cys	Asp	Glu	Cys	His	Ser	Thr	
		115					120					125				
GAT	GCC	ACA	TCC	ATC	TTG	GGC	ATC	GGC	ACT	GTC	CTT	GAC	CAA	GGA	GAG	432
Asp	Ala	Thr	Ser	Ile	Leu	Gly	Ile	Gly	Thr	Val	Leu	Asp	Gln	Gly	Glu	
	130					135					140					

ACT	GCG	GGG	GCG	AAA	TTG	GTT	GTG	TTC	GCC	ACC	GCC	ACC	CCT	CCG	GGC	480
Thr	Ala	Gly	Ala	Lys	Leu	Val	Val	Phe	Ala	Thr	Ala	Thr	Pro	Pro	Gly	
145					150				155						160	
TCC	GTC	ACT	GTG	CCC	CAT	CCC	AAC	ATT	GAG	GAG	GTT	GCT	CTA	TCC	ACC	528
Ser	Val	Thr	Val	Pro	His	Pro	Asn	Ile	Glu	Glu	Val	Ala	Leu	Ser	Thr	
				165					170					175		
ACC	GGA	GAG	ATC	CCT	TTT	TAC	GGC	AAG	GCT	ATC	CCC	CTT	GAG	GTA	ATC	576
Thr	Gly	Glu	Ile	Pro	Phe	Tyr	Gly	Lys	Ala	Ile	Pro	Leu	Glu	Val	Ile	
			180					185					190			
AAG	GGG	GGG	AGA	CAT	CTC	ATC	TTC	TGT	CAT	TCA	AAG	AGG	AAG	TGC	GAT	624
Lys	Gly	Gly	Arg	His	Leu	Ile	Phe	Cys	His	Ser	Lys	Arg	Lys	Cys	Asp	
		195					200					205				
GAG	CTC	GCC	ACA	AAG	CTG	GTC	GCA	ATG	GGC	ATC	AAT	GCC	GTG	GCC	TAC	672
Glu	Leu	Ala	Thr	Lys	Leu	Val	Ala	Met	Gly	Ile	Asn	Ala	Val	Ala	Tyr	
	210					215					220					
TAC	CGC	GGT	CTT	GAC	GTG	TCC	GTC	ATC	CCG	ACC	AGC	GGT	GAT	GTT	GTC	720
Tyr	Arg	Gly	Leu	Asp	Val	Ser	Val	Ile	Pro	Thr	Ser	Gly	Asp	Val	Val	
225					230				235					240		
GTC	GTG	GCA	ACC	GAC	GCC	CTC	ATG	ACC	GGC	TAT	ACC	GGC	GAC	TTC	GAC	768
Val	Val	Ala	Thr	Asp	Ala	Leu	Met	Thr	Gly	Tyr	Thr	Gly	Asp	Phe	Asp	
				245					250					255		
TGG	GTG	ATA	GAC	TGC	AAC	ACG	TGT	GTC	ACT	CAG	ACA	GTC	GAT	TTC	AGC	816
Ser	Val	Ile	Asp	Cys	Asn	Thr	Cys	Val	Thr	Gln	Thr	Val	Asp	Phe	Ser	
			260					265					270			
CTT	GAC	CCT	ACC	TTC	ACC	ATT	GAG	ACG	ACC	ACA	CTT	CCC	CAG	GAT	GCT	864
Leu	Asp	Pro	Thr	Phe	Thr	Ile	Glu	Thr	Thr	Thr	Leu	Pro	Gln	Asp	Ala	
		275					280					285				
GTC	TCC	CGC	ACT	CAA	CGA	CGG										885
Val	Ser	Arg	Thr	Gln	Arg	Arg										
	290					295										

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 295 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Thr Met Ile Thr Asn Ser Arg Gly Ser Ile Met Lys Ser Pro Val

[illegible]

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
- (B) TYPE: nucleic acid
- (C) STRAND FORM: single
- (D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

AAGGGATCCA TCATGAAATC CCCGGTGTTT ACGGATAACT

40

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRAND FORM: single
- (D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4

GGGAAGCCTT AATTCTTACC GTCGTTGAGT GCGGGAGAC

39

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRAND FORM: single

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5

GAGGGATCCA TCATGAAAGC GGTGGACTTT ATCCCTGTG

39

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 33 base pairs

(B) TYPE: nucleic acid

(C) STRAND FORM: single

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6

GAGAAGCTTT TAACACGTGT TGCAGTCTAT CAC

33

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 39 base pairs

(B) TYPE: nucleic acid

(C) STRAND FORM: single

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: cDNA

106230"2E095850

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GAGGGATCCA TCATGAAACA CCTGCATGCT CCCACCGGC

39

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 33 base pairs

(B) TYPE: nucleic acid

(C) STRAND FORM: single

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8

GAGAAGCTTT TAATACCAAG CACAGCCTGC GTC

33

TCATGAAACA